



Automatic inspection of a residual resist layer by means of self-organizing map

Zaki Sabit Fawzi Philippe, Stéphane Robert, Bernard Bayard

► To cite this version:

Zaki Sabit Fawzi Philippe, Stéphane Robert, Bernard Bayard. Automatic inspection of a residual resist layer by means of self-organizing map. Optical Engineering, 2016, 55 (5), 10.1117/1.OE.55.5.054106 . ujm-01385446

HAL Id: ujm-01385446

<https://hal-ujm.archives-ouvertes.fr/ujm-01385446>

Submitted on 21 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Automatic inspection of a residual resist layer by means of self-organizing map

Zaki Sabit Fawzi Philippe
Stéphane Robert
Bernard Bayard

SPIE.

Zaki Sabit Fawzi Philippe, Stéphane Robert, Bernard Bayard, "Automatic inspection of a residual resist layer by means of self-organizing map," *Opt. Eng.* **55**(5), 054106 (2016), doi: 10.1117/1.OE.55.5.054106.

Automatic inspection of a residual resist layer by means of self-organizing map

Zaki Sabit Fawzi Philippe,^{a,b,*} Stéphane Robert,^a and Bernard Bayard^a

^aLaboratoire Hubert Curien—Centre National de la Recherche Scientifique Unité Mixte de Recherche 5516, Batiment F, 18 Rue Professeur Benoît Lauras, Saint Etienne 42000, France

^bLaboratoire des Technologies de l'Information et de la Communication—Institut Universitaire des Sciences et Techniques d'Abéché, BP 6077 Ndjamen, Chad

Abstract. Photolithography allows large-scale fabrication of nanocomponents in the semiconductor industry. This technique consists of manufacturing a desired pattern on a photoresist film transferred onto the substrate during the etching process. Therefore, the mask quality is essential for reliable etching. For example, the presence of a residual layer of resist might be considered as a mask defect and can lead to the failure of the etching process. We propose the use of a Kohonen self-organizing map for automatic detection of a residual layer from an ellipsometric signature. The feasibility of the suggested inspection by the use of a classification technique is discussed and simulations are carried out on a 750-nm period grating. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.OE.55.5.054106](https://doi.org/10.1117/1.OE.55.5.054106)]

Keywords: optical inspection; pattern recognition; neural networks; nondestructive testing; ellipsometry; metrology.

Paper 151143 received Aug. 19, 2015; accepted for publication Apr. 29, 2016; published online May 26, 2016.

1 Introduction

In recent years, developments in microelectronic industries have been completely guided by miniaturization, which allows easy integration of components in embedded systems. To keep track of this technological evolution, the large-scale manufacture of nanocomponents needs appropriate, fast, and reliable techniques of fabrication and characterization. Photolithography is the most usual technique employed in the manufacturing process. It consists of depositing a thin photoresist film on a substrate. The expected pattern is then transferred onto the resist film. The film is hence structured and acts as a mask for the substrate in the subsequent etching process. Several techniques can then be used, such as wet¹ or dry etching (reactive ion beam etching, reactive ion etching, or ion beam etching).^{2,3} The presence of a residual layer in the region that should be etched can be considered as a mask defect and must be inspected. Standard characterization techniques mainly used for dimensional metrology, such as atomic force microscopy and scanning electron microscopy, are not suitable for rapid, effective monitoring and are destructive in most cases. Scatterometry is an alternative method widely studied in scientific literature.^{4,5} This nondestructive technique is based on the analysis of an optical signature of the light scattered by a periodic patterned structure. Furthermore, it is adapted for the in-line characterization process.^{6,7} This technique involves an inverse problem resolution, which can be performed by a different approaches, such as classic optimization,⁸ metaheuristics⁹ methods, regression, or library searching methods.¹⁰ Artificial neural networks (ANNs) have also been introduced as multilayer perceptrons (MLP).¹¹ Since 2002, our laboratory has developed an expertise in using ANN by the use of an MLP in a scatterometric process.^{12–14} The resolution of the inverse scattering problem involves suitable hypotheses, such as

the geometrical shape of the sample. The identification of the reliable geometrical profile is critical. In 2008, we demonstrated the possibility of identifying profile geometry by using a classical MLP from the ellipsometric signature¹⁵ used in the characterization process. Other works¹⁶ have employed another classification method, such as support vector machines, combined with the library search to solve the inverse problem. In terms of results, this classification method shows performances similar or complementary to the ANN method.¹⁷ Our aim is to set up a full ANN characterization method to meet rapid in-line control requirements. In this paper, we expose the potentiality of the Kohonen self-organizing map (SOM)^{18–20} in a characterization process devoted to classifying the optical signatures used in classical scatterometry depending on their corresponding geometrical structure. Indeed, this kind of ANN has been particularly effective in various pattern recognition tasks involving very noisy signals.^{21–23} In this work, we use the SOM as a data analysis tool to analyze and image the presence of different classes in the measurements featuring the structure to be tested. Our priority is to validate the use of the suitable optical signature in the inverse problem resolution to detect the reliable geometrical model. Our second aim is to detect the optimal model operating with a reduced number of parameters in the next characterization process. To validate the process, we propose in this work to study a very simple case, the detection of residual layer in the specific context of photolithography. We demonstrate the feasibility of the classification discerning geometrical profiles including or excluding a residual layer in the context of scatterometry.

2 Profile Definition and Direct Problem

The assumed shape of the geometrical grating is the symmetric trapezoidal profile commonly used in the domain of scatterometry. The grating period is constant and fixed by the

*Address all correspondence to: Zaki Sabit Fawzi Philippe, E-mail: sabit.zaki@univ-st-etienne.fr

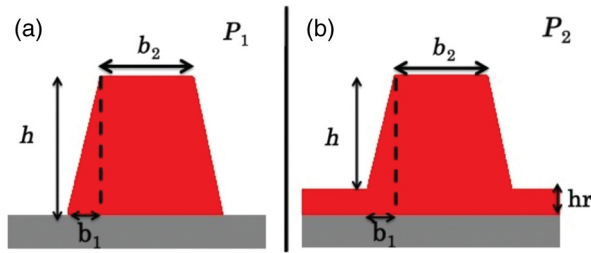


Fig. 1 Parametric description of two geometrical profiles: (a) P_1 is defined with three parameters: the sidewall projection b_1 , line width b_2 , and height h , and (b) P_2 presents an additional residual layer (thickness hr).

etching conditions. The classical profile P_1 is defined by three parameters: the sidewall projection b_1 , the line width b_2 , and the groove depth h [Fig. 1(a)]. A fourth parameter hr is merely added in the second profile P_2 to take into account the thickness of the residual layer [Fig. 1(b)].

In ellipsometric scatterometry, the optical signature is usually defined as measured intensities depending on the well-known ellipsometric parameters Ψ and Δ of the structure under testing. In the case of the phase-modulated ellipsometer used in this work, it concerns the intensities I_s and I_c for each wavelength.

$$I_s = \sin 2\psi \sin \Delta, \quad (1)$$

$$I_c = \sin 2\psi \cos \Delta. \quad (2)$$

Simulated signatures for a given profile can also be easily calculated by the multilayer modal method by Fourier expansion.²⁴⁻²⁶ In our work, a large set of simulated signatures are supplied to an SOM during the training step to identify them with the corresponding classes given by the profiles P_1 and P_2 .

3 Self-Organizing Map

Introduced by Kohonen,²⁷ SOM is one of the most popular ANN that is suitable and efficient for performing unsupervised clustering. Basically, the SOM can display a high-dimensional input space onto a reduced one to ensure a graphical visualization of classes. It is a technique similar to the principal component analysis²⁸ (PCA) method, which performs dimensional reduction. The difference between the two approaches is that the PCA is a linear mapping technique while the SOM performs a nonlinear lower-dimensional mapping. It has been successfully applied in engineering applications, such as pattern recognition, process control, or fault detection.

The SOM consists of S processing units (neurons), which can be arranged linearly, on a two-dimensional (2-D) grid, or in a three-dimensional (3-D) volume. Figure 2 shows a 2-D SOM network structure that is used in this work.

For a given SOM, the input vector \mathbf{X} has a fixed dimension n . The n elements of the vector \mathbf{X} (i.e., x_1, x_2, \dots, x_n) are connected to each neuron in the map and represent in our case the optical signature. A synaptic weight w_{ji} is defined for a connection from the i 'th component of the input vector to the j 'th neuron. Thus, each neuron j is represented by an n -dimensional weight vector $\mathbf{w}_j = [w_{j1}, \dots, w_{jn}]^T$ in the input space of the signatures. Neurons are also connected

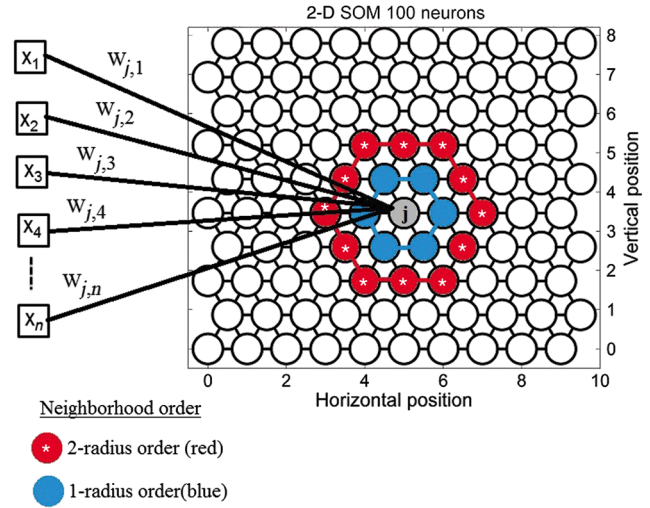


Fig. 2 Structure of a 2-D SOM defined by 10×10 neurons with n inputs and neighborhood size [blue first order and red (with*) second order] of considered neuron j .

to their neighbors. This particular connection is weighted by a neighborhood function taking into account the neurons topology during the calculation of the activation of each neuron. The neighborhood of a neuron is set by all neurons included in a circle, square, or hexagon focused on the considered neuron. Different orders of neighborhood can be defined according to the distance from the neuron. Adjacent neurons of a neuron j belong to the 1-neighborhood size. Each neuron is located by its horizontal and vertical position in the 2-D SOM. Figure 2 shows the concept of neighborhood size concerning a particular neuron j in the case of regular hexagon 2-D grid of neurons used in the work.

The number of neurons is defined before the training process. It determines the granularity of the mapping that affects the accuracy and generalization capacity of the map. The SOM training algorithm is the process devoted to preserving the topology of the high-dimensional input space onto the reduced output mapped space. Hence, the relative distances between data points are preserved. During the iterative training, the SOM determines the similarity of the input vector \mathbf{X} and neurons represented by the vector \mathbf{w}_j in terms of Euclidian distance. The neuron whose weight vector is closest to the considered input \mathbf{X} is considered as winner and called best matching unit (BMU). It is denoted by c in the following expression:

$$\|\mathbf{x} - \mathbf{w}_c\| = \min_j \|\mathbf{x} - \mathbf{w}_j\|. \quad (3)$$

Once a BMU is found, its associated weight and those of its neighbors are updated in accordance with the following Kohonen rule in the next iteration:

$$\mathbf{w}_j(t+1) = \mathbf{w}_j(t) + \alpha(t) \times h_{j,c}(t) [\mathbf{x}(t) - \mathbf{w}_j(t)], \quad (4)$$

where $\alpha(t)$ and $h_{j,c}(t)$ are, respectively, the learning rate and the neighborhood function at iteration t . $\alpha(t)$ decreases from α_0 during the iteration t and is close within 0 and 1.

$$\alpha(t) = \frac{\alpha_0}{1+t}. \quad (5)$$

The neighborhood function $h_{ic}(t)$ depends on the distance between each neuron j and the BMU. We use the particular Gaussian kernel function defined by the parameter σ by

$$h_{j,c} = \exp\left(-\frac{r_{j,c}^2}{2\sigma_t^2}\right), \quad (6)$$

where $r_{j,c} = \|r_j - r_c\|$ is the distance from the position r_c of the BMU c to the position r_j of the neighborhood neuron j on the map. σ_t represents the neighborhood radius at iteration t . As with the learning rate, σ_t is reduced during the training and converges to the BMU (Fig. 2). The initial value depends on the map size and the final value is close to 1. Initial weights are randomly chosen from the range $[0, 1]$ and neighborhood initial value is fixed at half the size of the map.

After a fixed number of iterations, the learning procedure leads to ordered mapping of the input data. Hence, similar patterns are supposed to be mapped in the same region on the map, while dissimilar ones are more dispersed.

4 Theoretical Results

The diffracting structure used for the test has a period of 750 nm. The grating is an NEB22 resist deposited on a silicon substrate. As previously mentioned, two types of profiles (P_1 and P_2) are considered. Geometrical parameters for 6000 theoretical samples (3000 for each profile) are randomly generated from the following ranges:

$$0 \leq b_1 \leq 70 \text{ nm}, \quad 100 \leq b_2 \leq 350 \text{ nm},$$

$$170 \leq h \leq 250 \text{ nm (profile } P_1)$$

$$0 \leq b_1 \leq 70 \text{ nm}, \quad 100 \leq b_2 \leq 350 \text{ nm},$$

$$170 \leq h \leq 250 \text{ nm}, \quad 1 \leq \text{hr} \leq 100 \text{ nm (profile } P_2)$$

Ellipsometric signatures $\{I_s, I_c\}$ are calculated for wavelengths between 1.5 and 6.5 eV with a step of 0.2 eV (52 values, i.e., 26 for each intensity) in the experimental condition (incidence angle of 70 deg). Gaussian noise with zero mean and 3% standard deviation is considered in order to take experimental conditions into account. Simulated signatures are linearly normalized and then supplied to an SOM composed of 400 neurons arranged in a 2-D grid.

The ANN is trained with the whole simulated database. At the end of the process, the clustering of the P_1 and P_2 profiles is observable. The SOM can hence respond to new request by a vector of 400 elements corresponding to each activation of neurons and can be plotted in a 2-D space (20×20). Figure 3 shows the distribution on the map of all P_1 profile (3000) after training. Each cell represents a neuron and the integrated number defines the number of similar profiles that activate the considered neuron. We notice that the P_1 profile stimulates specific neurons (blue in the figure) and they are distributed in a very distinctive area in the SOM space.

The same kind of consideration can be made with the profile P_2 (Fig. 4). Most P_2 profiles (2607 samples from 3000) also stimulate specific neurons (in green), but a small

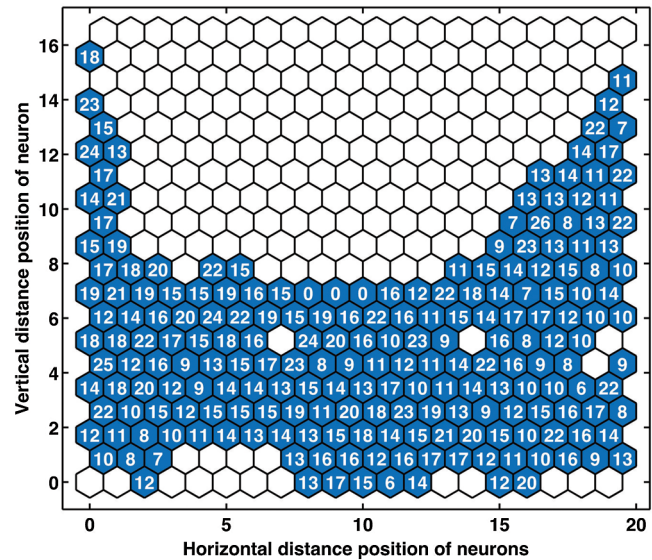


Fig. 3 Results of SOM (20×20 neurons) concerning the distribution of profiles P_1 from ellipsometric signatures (3000 samples composing the training data).

proportion of them (393 samples, i.e., 13.1% of P_2 profiles) stimulates neurons that have already been activated mostly by the P_1 profile (blue with '*'). The latter profiles are characterized by a small residual layer thickness h_r under 15 nm.

To complete this study, we proceed to a labeling of the map neurons. We define the probability of a specific neuron j to be sensitive to a fixed profile P_i ($i = 1, 2$) by

$$p_j(P_2) = \frac{k}{m}, \quad (7)$$

where m is the number of samples activating the specific neuron j and k is the number of P_i profiles.

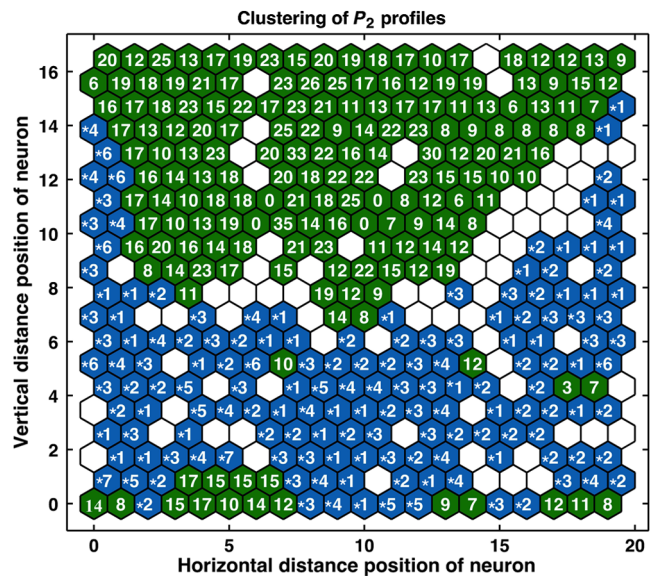


Fig. 4 Results of SOM (20×20 neurons) concerning the distribution of profiles P_2 from ellipsometric signatures (3000 samples composing the training data). Green neurons are entirely activated by P_2 profile and blue neurons (with '*') are those activated by both P_2 and P_1 profiles (Fig. 3).

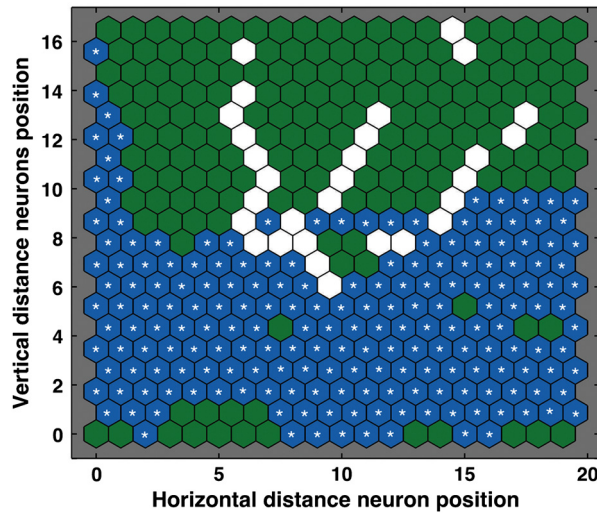


Fig. 5 SOM labeled for the classification. Green neurons are sensible to P_2 , blue neurons (with*) are sensible to P_1 , and white neurons are never activated.

The probability $p_j(P_2)$ of all neurons shaded in blue in Fig. 3 is very low or close to zero. A limit level is defined in order to make a decision regarding class by fixing a probability $p_j(P_2)$. A simple choice can be made with a value of 0.5. Hence, a signature activating a neuron j with a $p_j(P_2)$ probability >0.5 will be considered as P_2 profile and otherwise P_1 profile. Figure 5 shows the corresponding labelled SOM. Green units match neurons sensitive to P_2 profile, blue units match neurons sensitive to P_1 , and white units have not been activated by any type of profile considered in the training set.

In order to estimate the quality of the inspection, a limit of detection between the two types of profile in the SOM space can be determined for a fixed sample. First, we consider three profiles with a significant thickness hr.

Sample 1: $b_1 = 0.5$ nm, $b_2 = 131.4$ nm, $h = 182.6$ nm, $hr = 31.6$ nm

Sample 2: $b_1 = 3.9$ nm, $b_2 = 103.3$ nm, $h = 238.1$ nm, $hr = 65.6$ nm

Sample 3: $b_1 = 58.1$ nm, $b_2 = 176.2$ nm, $h = 220.2$ nm, $hr = 23.3$ nm

For each case, a set of samples is then simulated by reducing the hr by a step of 2 nm until 0. All the corresponding signatures are generated. The signatures are supplied to the previous trained and labeled SOM. The probability $p_{\text{BMU}}(P_2)$ of BMU is shown in Fig. 6 for each set of samples. As regards Fig. 6, the limit of detection is estimated between 8 and 17 nm. This depends on the considered sample and, by extension, on the neuron sensitivity.

In order to improve this limit, we construct an SOM with an additional layer of neurons (3-D grid of neurons). The SOM is made up of 12 layers of 12×12 neurons (1728 neurons). Only 7.2% of P_2 profiles (216 samples instead of 393 in Fig. 4) stimulate neurons preferentially activated by P_1 . The previous study is then performed in the same conditions. Signatures of the three sets of samples defined above are then supplied to the 3-D SOM. The results are shown in

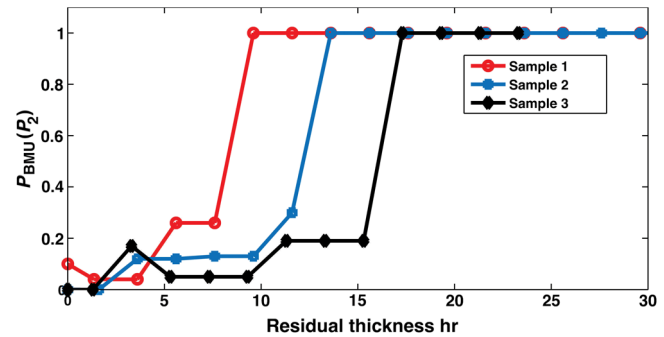


Fig. 6 $p_{\text{BMU}}(P_2)$ probability of the activated neuron from the 2-D SOM (20×20 neurons) versus the corresponding theoretical hr for the three specific samples.

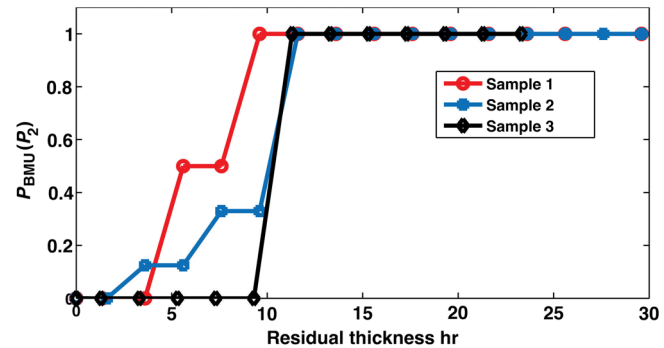


Fig. 7 $p_{\text{BMU}}(P_2)$ membership probability of the activated neuron from a 3-D SOM ($12 \times 12 \times 12$ neurons) versus the corresponding theoretical hr for the three specific samples.

Fig. 7. In this case, the limit of detection is now between 4 and 11.5 nm.

3-D SOM significantly improves the results and displays better performances. The maximum detection limit of the residual thickness is reduced to around 11.5 nm instead of 17 nm for the 2-D SOM. We note that some neurons showed result neurons down to 4 nm (sample 1). The value of the detection limit depends on the samples under testing and on the architecture of the SOM.

5 Conclusion

In this paper, we have presented the potentiality of the ANN to classify grating shape from a classical optical signature used in scatterometry. An SOM has been developed to extract qualitative information on the geometrical profile shape. An estimation of the presence of a mask defect in lithography can also be calculated to assist in decision making. Simulated results have been carried out. The limit of the detection has been simply estimated in three specific cases. It is also important to notice that the analysis proposed does not supply a binary decision, such as with a classical method, but estimates a probability for each model, which can be considered as a confidence index. This is more reliable for making the best decision.

In the work, we emphasize the importance in verifying that optical measurements used in the characterization process include qualitative information on the geometrical model. We have theoretically demonstrated that the SOM is a practical tool to get a rapid visualization of the capacity of the

optical signature to classify different geometrical models. The technique has the advantage of being easily applied to identification of various kinds of defects with reduced prior knowledge of the scatterometric profile. Next, a complete experimental study must be undertaken for accurate validation of the technique. An effective classification method must be applied. A supervised classifier, such as learning vector quantization, may be used to take advantage of the SOM analysis. We will concentrate our future efforts on these points in order to improve the detectable limit by optimizing the ANN. Looking further ahead, this promising path can lead us to ignore very restrictive structural hypotheses in a scatterometry process.

Acknowledgments

The authors are grateful to P. Donnelly for stimulating discussions leading to a more comprehensive presentation of this work.

References

1. J. W. Kim et al., "Fabrication of inverse micro/nano pyramid structures using soft UV-NIL and wet chemical methods for residual layer removal and Si-etching," *Microelectron. Eng.* **110**, 403–407 (2013).
2. J. M. E. Harper, "Ion beam etching," in *Plasma Etching-An Introduction*, D. M. Manos and D. L. Flamm, Eds., pp. 391–423, Academic Press, San Diego (1989).
3. J. W. Coburn, H. F. Winters, and T. J. Chuang, "Surface processes in plasma-assisted etching environments," *J. Vac. Sci. Technol. B* **1**, 496 (1983).
4. I. Gereige et al., "Dimensional characterization of biperiodic imprinted structures using optical scatterometry," *Microelectron. Eng.* **112**, 27–30 (2013).
5. C. J. Raymond et al., "Multiparameter grating metrology using optical scatterometry," *J. Vac. Sci. Technol. B* **15**, 361 (1997).
6. M. El kodadi et al., "Resist trimming etch process control using dynamic scatterometry," *Microelectron. Eng.* **86**, 1040–1042 (2009).
7. S. Soulan et al., "In line etching process control using dynamic scatterometry," *Proc. SPIE* **6617**, 661713 (2007).
8. C. J. Raymond et al., "Comparison of solutions to the scatterometry inverse problem," *Proc. SPIE* **5375**, 564–575 (2004).
9. G. Cormier and R. Boudreau, "Genetic algorithm for ellipsometric data inversion of absorbing layers," *J. Opt. Soc. Am. A* **17**, 129–134 (2000).
10. X. Chena et al., "Improved measurement accuracy in optical scatterometry using fitting error interpolation based library search," *Measurement* **46**(8), 2638–2646 (2013).
11. I. J. Kallioniemi and J. Saarinen, "Optical scatterometry with neural network model for nondestructive measurement of submicron features," *Proc. SPIE* **3743**, 33–40 (1999).
12. S. Robert, A. Mure-Ravaud, and D. Lacour, "Characterization of optical diffraction grating by use of a neural method," *J. Opt. Soc. Am. A* **19**(1), 24–32 (2002).
13. I. Gereige et al., "Rapid control of submicrometer periodic structures by a neural inversion from ellipsometric measurement," *Opt. Commun.* **278**, 270–273 (2007).
14. S. Robert and A. Mure-Ravaud, "Control of the homogeneity of an optical grating by a neural characterization," *Opt. Eng.* **44**(3), 033601 (2005).
15. I. Gereige, S. Robert, and J. Eid, "Automatic detection of photoresist residual layer in lithography using a neural classification approach," *Microelectron. Eng.* **97**, 29–32 (2012).
16. J. Zhu et al., "Identification and reconstruction of diffraction structures in optical scatterometry using support vector machine method," *J. Micro/Nanolithogr. MEMS MOEMS* **12**(1), 013004 (2013).
17. J. Ren, "ANN vs. SVM: which one performs better in classification of MCCs in mammogram imaging," *Knowl. Based Syst.* **26**, 144–153 (2012).
18. T. Kohonen, "The self organising map," *Neurocomputing* **21**(1–3), 1–6 (1998).
19. A. Mekler and D. Schwarz, "Quality assessment of data discrimination using self-organizing maps," *J. Biomed. Inf.* **51**, 210–218 (2014).
20. F. Coleca et al., "Self-organizing maps for hand and full body tracking," *Neurocomputing* **147**, 174–184 (2015).
21. A. Majumder, L. Behera, and V. K. Subramanian, "Emotion recognition from geometric facial features using self-organizing map," *Pattern Recognit.* **47**(3), 1282–1293 (2014).
22. M. S. Prieto and A. R. Allen, "Using self-organising maps in the detection and recognition of road signs," *Image Vis. Comput.* **27**(6), 673–683 (2009).
23. E. Alhoniemi et al., "Process monitoring and modeling using the self organising map," *Integr. Comput. Aided Eng.* **6**(1), 3–14 (1999).
24. R. Anderwartha, G. H. Derrick, and R. C. McPhedran, "A general model theory for reflection gratings," *Opt. Acta* **28**, 1501–1516 (1981).
25. P. Schiavone, G. Granet, and J. Y. Robic, "Rigorous electromagnetic simulation of EUV masks: influence of the absorber properties," *Microelectron. Eng.* **57–58**, 497–503 (2001).
26. L. Li, "New formulation of the Fourier modal method for crossed surface-relief gratings," *J. Opt. Soc. Am. A* **14**(10), 2758–2767 (1997).
27. T. Kohonen, "The self organising map," *Proc. IEEE* **78**(9), 1464–1480 (1990).
28. I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New York (1986).

Zaki Sabit Fawzi Philippe received his master's degree in electrical engineering from the Institute University of Science and Technology of Abeché, Chad, in 2011. Currently, he is a PhD candidate under the guidance of Bernard Bayard and Stéphane Robert in the Laboratory Hubert Curien of the University of Saint-Etienne, France. His research interests include recognition and neural classification of patterned optical structures including modelling in noncontact characterization methods (scatterometry and ellipsometry).

Stéphane Robert received his PhD in the field of optoelectronics from the University of Saint-Etienne in 2003. Currently, he is working in the Laboratory Hubert Curien, especially on polarized light and magneto-optic properties of materials. His research interests include noncontact characterization methods (scatterometry and ellipsometry) of optical structures and the development of the associated optimization process (neural network and genetic algorithm).

Bernard Bayard received his master's degree and PhD in electronics from the University of Saint-Etienne, France, in 1997 and 2000, respectively. Currently, he is an assistant professor at the University of Saint-Etienne, France. His research activities concern the integration of magnetic microwave components and the microwave characterization of materials.